

1. Xây dựng kế hoạch mang tính chiến lược cho việc sử dụng GIS
2. Tổ chức cơ sở dữ liệu trong GIS
3. CPU Structure and Functions

Xây dựng kế hoạch mang tính chiến lược cho việc sử dụng GIS  
Xây dựng kế hoạch mang tính chiến lược cho việc sử dụng GIS

## **NHỮNG VẤN ĐỀ CẦN QUAN TÂM TRONG TỔ CHỨC THỰC HIỆN HỆ THỐNG GIS**

### **Tổ chức hệ thống thông tin địa lý**

Tổ chức hệ thống thông tin địa lý phải bắt đầu bằng nhiệm vụ đặt ra của hệ thống thông tin. Từ đó xác định được mức độ đòi hỏi của các loại thông tin cần thiết như thông tin thuộc tính tỷ lệ nào, phải có lớp thông tin nào, độ chính xác của thông tin và thông tin thuộc tính có dạng nào. Sau khi xác định được nhu cầu thông tin cần tìm xem thông tin này có được từ nguồn nào, có thể lấy được từ cơ sở dữ liệu quốc gia, cơ sở dữ liệu chuyên ngành mua trên thị trường thông tin. Nếu các nguồn khai thác thông tin đều chưa có hoặc có cần xác định tiếp sử dụng biện pháp nào để thu nhận các thông tin còn thiếu. Có thể được các công ty cung cấp thông tin hoặc tổ chức thu nhập hệ thống thông tin riêng cho số lượng dữ liệu cần có nên tổ chức quản lý dưới dạng tập trung hay phân tán cho nhu cầu và hoàn cảnh khai thác thông tin. Khi định dạng thông tin đã rõ mới thiết kế phần cứng và phần mềm phù hợp. Không cần có tham số mạnh hơn nhu cầu đòi hỏi cũng không yếu hơn để không thực hiện được nhiệm vụ cần thực hiện. Nếu cơ sở dữ liệu định dạng là phân tán thì phải thiết kế phần cứng dưới dạng máy tính (intranet). Khi dữ liệu cho phép các hệ thống thông tin khác truy nhập đến cần thiết kế các cổng extranet. Từ định hình cơ sở dữ liệu, nhu cầu quản lý và phần cứng đã được xác định tiếp tục xác định các phần mềm hệ thống, phần mềm quản trị dữ liệu địa lý và thuộc tính phần mềm quản trị nếu cần, phần mềm ứng dụng. Đôi khi các nhà tổ chức HTTT xét lựa chọn một giải pháp công nghệ hợp lý.

### **Cơ sở dữ liệu địa lý và cơ sở dữ liệu thuộc tính**

Phần mềm đã giới thiệu cơ sở dữ liệu của hệ thống thông tin địa lý có hai thành phần: cơ sở dữ liệu địa lý và cơ sở dữ liệu thuộc tính. Về lý thuyết có thể tổ chức quản lý cả 2 loại dữ liệu có cùng một cơ sở dữ liệu và do cùng một hệ quản trị dữ liệu. Lúc này các dữ liệu thông tin được coi như các lớp riêng biệt gắn liền với các đối tượng địa lý. Trong thực tế vấn đề trở nên phức tạp khi số lượng dữ liệu quá lớn, vấn đề tìm thông tin và truy cập thông tin được giải quyết. Vì vậy người ta tổ chức quản lý các dữ liệu này trong hai cơ sở dữ liệu riêng biệt với hai hệ quản trị riêng biệt.

Cơ sở dữ liệu thuộc tính được lưu trữ tương tự như các loại cơ sở dữ liệu thông thường: ngân hàng, luật pháp, hành chính, v.v.. các dữ liệu được lưu trữ dưới dạng các cột gọi là trường (field) và các hàng gọi là bản tin (record). Vấn đề sẽ không bàn luận gì nhiều nếu có một bản. Thực tế bức tranh quan hệ giữa các dữ liệu đã làm phải có tổ chức quản lý ở nhiều bản khác nhau và phải tìm được một số trường mô tả được mối quan hệ giữa các bản. Căn cứ vào các mối quan hệ này để giải quyết việc tìm dữ liệu và cập nhật dữ liệu. Dạng tổ chức cơ sở dữ liệu như vậy được gọi là cơ sở dữ liệu quan hệ. Hiện nay quản trị cơ sở dữ liệu thường gặp đều ở dạng cơ sở dữ liệu quan hệ như DBase, INFOMI, SQL Serve, ORACLE..v.v. Sự khác nhau giữa các hệ quản trị dữ liệu là dữ liệu (format) và ngôn ngữ hỏi đáp để tìm dữ liệu và cập nhật dữ liệu.

Cơ sở dữ liệu địa lý (chủ yếu là bản đồ) không thể lưu trữ và quản lý theo nguyên tắc các loại cơ sở dữ liệu thông thường nói trên. Các dữ liệu bản đồ được cấu tạo bởi các thành phần là tọa độ các điểm, dãy tọa độ các điểm nối giữa 2 điểm (đường) và dãy cùng nối liên tiếp trong một nhóm điểm (miền). Như vậy bức tranh dữ liệu ở đây khác so với dữ liệu thông thường. Vì chúng ta có các đối tượng địa lý khác nhau nên các điểm, đường nét của từng loại đối tượng địa lý sẽ được quản lý riêng trong từng lớp thông tin. Trong thông tin này có một lớp thông tin cơ bản nhất là lớp về hệ quy chiếu (lưới tọa độ) để thể hiện các lớp thông tin theo yêu cầu. Hệ quản trị cơ sở dữ liệu địa lý quen thuộc là ARC/INFO và MGE. Mỗi hệ cũng có định dạng dữ liệu khác nhau, xác định hình học khác nhau, xác định quan hệ hình học khác nhau và thủ tục tìm - cập nhật khác nhau.

## Quản trị dữ liệu và khai thác dữ liệu

Hai hệ cơ sở dữ liệu thuộc tính và địa lý nói trên đã được giải quyết khá chọn lọc trong việc quản trị từng cơ sở dữ liệu. Công việc quản trị dữ liệu bao gồm nhiều công việc như:

- Đưa thông tin vào cơ sở dữ liệu
- Tìm nhanh chóng các dữ liệu cần thiết
- Tổ chức cập nhật dữ liệu mới gồm việc xoá bỏ các dữ liệu không cần bổ sung các dữ liệu mới và hiệu chỉnh các dữ liệu đang có (hãy lưu ý rằng việc khó khăn tiếp tục cập nhật là do mối quan hệ phức tạp giữa các dữ liệu).
- Phân chia các nhóm thông tin theo phân cấp bảo mật thông tin:
- Đảm bảo an toàn thông tin trong cơ sở dữ liệu (không mất thông tin ngay trường hợp hỏng hóc toàn hệ thống).
- Đánh giá hiện trạng của dữ liệu
- Xuất thông tin theo nhu cầu của người sử dụng.

Trong hệ thống thông tin địa lý, vấn đề quản trị dữ liệu có một đặc thù riêng. Đó là vấn đề khai thác liên kết giữa các cơ sở địa lý và cơ sở dữ liệu thuộc tính. Vấn đề đặt ra là tìm mối quan hệ giữa đối tượng địa lý trong cơ sở dữ liệu địa lý với thuộc tính của nó trong cơ sở dữ liệu thuộc tính và ngược lại. Lý thuyết thông tin hiện đại đã giải quyết trọn vẹn việc quyết định quan hệ này. Cách giải quyết có thể đi theo hướng nhúng dữ liệu của cơ sở dữ liệu này vào môi trường của cơ sở dữ liệu kia hoặc có thể tìm một môi trường trung gian để thể hiện loại dữ kiện.

Khai thác dữ kiện là nhu cầu đặt ra có tầm quan trọng đặc biệt để giúp áp dụng mọi điều kiện thuận lợi tiếp cận với dữ liệu trong phạm vi được phép khai thác được thể hiện được các dạng sau:

- Tìm các nhóm dữ liệu để cung cấp cho các cơ sở dữ liệu khác thông qua đặc tính hoặc các thiết bị ngân hàng.
- Phân tích tổng hợp thông tin để trả lời một số câu hỏi mang tính quản lý.

- Thực hiện các bài tính Ứng dụng cho nhiều mục đích khác nhau như quản lý sản xuất kinh doanh, quản lý xã hội, nghiên cứu khoa học, giáo dục đào tạo, nâng cao dân trí.
- Hiện thị dữ liệu ở các dạng khác nhau như bản đồ, biểu đồ, đồ thị, bản số, bản phục vụ các nhu cầu kinh tế - xã hội

## **Định chuẩn hệ thống và hệ thống mở**

Hệ thống thông tin càng phát huy tác dụng rộng rãi nếu hệ thống này được định. Điều kiện này giúp cho thông tin được trao đổi rộng rãi hơn, người sử dụng thông hơn, hệ thống cập nhật thông tin đa dạng hơn, hiệu quả là tránh được lãng phí trong phát triển. Muốn vậy các ngành cần thống nhất một chuẩn chung mang tính Quốc gia định chuẩn có thể hiểu một cách đơn giản là mọi hệ thống thông tin đều có thể hiện thị của nhau. Vấn đề chuẩn hoá thông tin đang là nội dung chủ yếu trong giai đoạn công nghệ thông tin hiện nay. Công việc chuẩn hoá thông tin bao gồm các nội dung sau:

- Chuẩn hoá môi trường tin học trong môi trường hệ thống, đồ hoạ cơ sở dữ liệu.
- Chuẩn hoá thông tin tiếng việt.
- Chuẩn hoá định dạng (FORMAT) thông tin ở các thể loại: chữ- số, vector, Multimedia, v.v.
- Chuẩn hoá hệ qui chiếu Quốc gia.
- Chuẩn hoá hệ tọa độ và cao độ nhà nước.
- Chuẩn hoá hệ thống địa danh và địa giới hành chính các cấp.
- Chuẩn hoá việc định nghĩa các đối tượng địa lý, nội dung bản đồ các ký hiệu, các phân lớp thông tin địa lý.
- Chuẩn hoá đường chuyển dữ liệu trên mạng.

Vấn đề chuẩn hoá cũng cần tránh xu hướng quá khích trong việc cưỡng bức một số sản phẩm phần cứng và phần mềm nào đó.

Vấn đề chuẩn hoá sẽ được giải quyết nhẹ nhàng hơn khi con người đưa ra khá nhiều hệ thống mở. Khái niệm này được thể hiện ở các khía cạnh

sau:

- Mở về phần cứng là đảm bảo không bị lạc hậu khi công nghệ phần cứng phát triển có thể liên kết với các phần cứng khác;
- Mở về phần mềm là việc công bố rõ các giao diện ở đầu vào và đầu ra cùng khả năng kết nối với các phần mềm ứng dụng khác;
- Mở về cơ sở dữ liệu có các giao diện chung với các cơ sở dữ liệu khác.
- Tính mở của hệ thống không chỉ để giải quyết việc giảm nhẹ công việc chuẩn còn tạo khả năng dễ dàng mở rộng hệ thống trong quá trình công nghệ thông tin phát triển mạnh mẽ như hiện nay.

## **Hệ thống thông tin địa lý và mạng vi tính - Internet**

Xây dựng một hệ thống thông tin địa lý cho các cơ sở dữ liệu nhỏ chỉ cần một hoặc một nhóm máy tính nối với nhau bằng mạng cục bộ. Vấn đề mạng máy tính trong trường hợp này không đặt ra nhiều lắm. Cấu trúc mạng có dạng một máy tính đóng vai trò máy chủ nối với một số máy tính đảm nhận nhiệm vụ dữ liệu, một số máy tính thực hiện việc ứng dụng vào một số máy tính hiển thị thông tin.

Đối với một cơ sở dữ liệu lớn người ta phải tổ chức các cơ sở dữ liệu phân tán thành nhiều cơ sở dữ liệu nhỏ nối với nhau bằng mạng diện rộng. Mỗi cơ sở dữ liệu nhỏ tổ chức như ở trên có gắn thêm một máy tính đóng vai trò netserver cùng với một bộ liệu đường dài gồm modem và router. Vấn đề cần giải quyết là đường truyền dữ liệu, an toàn dữ liệu trên đường truyền và bảo mật dữ liệu.

## **Các vấn đề chi phí tổ chức hệ thống GIS**

Ở các nước đang và kém phát triển, chi phí đầu tư khá lớn để tổ chức và duy trì hoạt động của một hệ thống GIS là trở ngại rất quan trọng trong quá trình phát triển công nghệ GIS. Chi phí sẽ bao gồm:

- Đầu tư ban đầu cho việc lắp đặt trang thiết bị và phần mềm trong đó chi phí dành cho cài đặt phần mềm GIS gấp 10 lần chi phí mua sắm thiết bị.
- Chi phí lớn để chuẩn bị và hoàn thiện các nguồn số liệu hiện có nhằm đáp ứng với tiêu chuẩn của hệ thống GIS. Tỷ lệ đầu tư cho số liệu là 100 lần so với 10 lần của phần mềm, Như vậy tỷ lệ đầu tư chung sẽ là:

Thiết bị GIS Phần mềm GIS Số liệu theo yêu cầu

(1 lần) (10 lần) (100 lần)

Trong quá trình hoạt động còn nảy sinh các chi phí tiềm ẩn khác (Hidden costs) để (i) mua sắm vật tư đặc chủng (ii) thay thế và nâng cấp khả năng thiết bị hay phần mềm (iii) lắp đặt thiết bị ngoại vi để lưu trữ có hiệu quả những khối lượng số liệu ngày càng lớn.

Ngoài ra, hiện nay với những nhu cầu sử dụng GIS ngày càng tăng kéo theo sự đổi mới nhanh chóng thiết bị và phần mềm, nguồn chi phí lớn để tái đầu tư, mở rộng và hiện đại hoá sẽ là một đòi hỏi với bất kỳ hệ thống GIS nào trong giai đoạn 3 đến 5 năm.

## **Những hợp phần thiết yếu cho sự hoạt động thành công hệ thống GIS**

Bên cạnh vấn đề thiết bị và phần mềm, một hệ thống GIS chỉ hoạt động thành công khi những hợp phần thiết yếu sau đây được quan tâm và đầu tư đúng mức, đó là:

- Nhân sự và huấn luyện

Một hệ thống GIS hoạt động có hiệu quả khi được điều hành bởi các chuyên gia có kinh nghiệm và những kỹ thuật viên GIS thành thạo. Để có đội ngũ nhân sự này, một kế hoạch huấn luyện gồm nhiều giai đoạn và nhiều trình độ khác nhau (trong nước và ngoài nước) cần được soạn thảo

và thực hiện chu đáo, kế hoạch này bao gồm cả các dạng huấn luyện trong công việc (on - job training) là mô hình phổ biến hiện nay để có thể chuyển giao được công nghệ một cách nhanh chóng, có hiệu quả và tiết kiệm.

- Bảo dưỡng thiết bị và hỗ trợ hoạt động GIS

Là dạng công nghệ mới nên vấn đề bảo dưỡng thiết bị GIS là yếu tố cần thiết để duy trì tốt hoạt động của hệ thống. Những trục trặc ở phần cứng (hardware) và phần mềm (software) cũng như ở các thiết bị ngoại vi, đòi hỏi phải được ngăn ngừa và thay thế bởi chế độ bảo dưỡng định kỳ do các chuyên gia có kinh nghiệm thực hiện. Bên cạnh đó, cần thiết lập được mạng lưới các đơn vị hay công ty tư vấn để hỗ trợ các hoạt động của hệ thống GIS, bao gồm (y) tư vấn giải pháp khi có sự cố (ii) tư vấn việc nâng cấp và thay thế hay mở rộng thiết bị, phần mềm (iii) cung ứng kịp thời vật tư và thiết bị thay thế v.v...

- Số liệu

Trong hoạt động GIS, số liệu là một trong những hợp phần thiết yếu nhất và được xem là nguồn "nhiên liệu" của hệ thống. Do đó, tổ chức thu thập số liệu hiện có và cung ứng số liệu thường xuyên là nội dung rất quan trọng để một hệ thống GIS có thể hoạt động được. Hiện nay, số liệu rất phân tán với nhiều nguồn và nhiều dạng khác nhau, nên tiến hành tái hiệu chỉnh số liệu hiện có theo một định dạng (format) tiêu chuẩn để có thể sử dụng trong GIS là công việc bắt buộc của hệ thống GIS.

- Tài chính

Đây là yếu tố mấu chốt của việc xây dựng và duy trì hoạt động của hệ thống GIS. Tài chính không những cần thiết cho việc đầu tư ban đầu và vận hành thiết bị - phần mềm GIS, các chương trình huấn luyện và chính sách đãi ngộ đối với lực lượng chuyên gia - kỹ thuật viên GIS cũng đòi hỏi những chi phí lớn. Do vậy, mức phát triển của hệ thống GIS cần căn cứ khả năng đáp ứng của nguồn tài chính, do nhiều chi phí sẽ phát sinh theo cấp số nhân so với đầu tư ban đầu về thiết bị. Tuy vậy, nguồn tài chính cũng sẽ được thu lại từ các hoạt động có hiệu quả của hệ thống



GIS. Bên cạnh đó, cần tính đến các lợi ích không thể tính bằng tiền do việc ứng dụng công nghệ GIS đem lại trong các lĩnh vực nghiên cứu - quản lý và quy hoạch phát triển.

- Sự phối hợp và cộng tác của những tổ chức cung cấp số liệu và sử dụng kết quả GIS

Một hệ thống GIS chỉ hoạt động có hiệu quả trong môi trường có sự cộng tác của nhiều tổ chức, cơ quan từ cấp trung ương đến địa phương. Có thể rằng "đầu vào" (input) và "đầu ra" (output) của một hệ thống GIS tùy thuộc vào yếu tố phối hợp và cộng tác của các đơn vị cung cấp số liệu và sử dụng kết quả chỉ với nguồn số liệu đầy đủ - phong phú và chính xác, hệ thống GIS mới hoạt động có hiệu quả và cung cấp các kết quả hữu ích, các kết quả này sẽ thuyết phục người sử dụng tiếp tục đặt ra các yêu cầu mới, từ đó kích thích và đẩy mạnh hoạt động của hệ thống GIS. Chu trình nói trên là quy luật tất yếu và là điều kiện cần có để một công nghệ như GIS có thể phát triển.

## **TIẾN TRÌNH TRIỂN KHAI MỘT HỆ THỐNG GIS**

Để triển khai một hệ thống GIS hoạt động có hiệu quả, một tiến trình bao gồm 4 vấn đề sau đây cần được xác định và giải quyết

### **Vấn đề tổ chức**

Hệ thống GIS sẽ được tổ chức như thế nào tại địa phương, bao gồm (i) quy mô quản lý của hệ thống và (ii) mô hình số liệu sẽ được sử dụng hay cung cấp từ hệ thống này,

### **Vấn đề kỹ thuật**

Cần xác định rõ được (i) thiết bị và phần mềm được cài đặt, (ii) trình độ và yêu cầu người sử dụng kết quả GIS và (iii) tổ chức hỗ trợ các hoạt

động GIS như thế nào.

### **Vấn đề số liệu**

Có 3 khía cạnh cần được quan tâm, đó là (i) khả năng thích ứng của các dạng số liệu đối với hệ thống GIS, (ii) tiêu chuẩn về chất lượng số liệu và (iii) tổ chức quản lý số liệu.

### **Vấn đề huấn luyện và nhân sự**

Cần chuẩn bị (i) một chương trình huấn luyện cơ bản và nâng cao, (ii) tổ chức bộ máy nhân sự và (iii) bố trí con người thích hợp.

## **KẾ HOẠCH MANG TÍNH CHIẾN LƯỢC TRONG ỨNG DỤNG CÔNG NGHỆ GIS**

Hiện nay GIS đã và đang được sử dụng ở các nước đang phát triển, và nó đã mang lại những kết quả khả quan. Tuy nhiên để có thể ứng dụng và sử dụng các trang thiết bị của GIS, điều quan trọng trước hết là cần phải xây dựng một kế hoạch đầu tư, sử dụng, bảo trì cũng như ứng dụng vào trong thực tế, tất cả các kế hoạch đó được xem như là kế hoạch mang tính chiến lược cho sự phát triển.

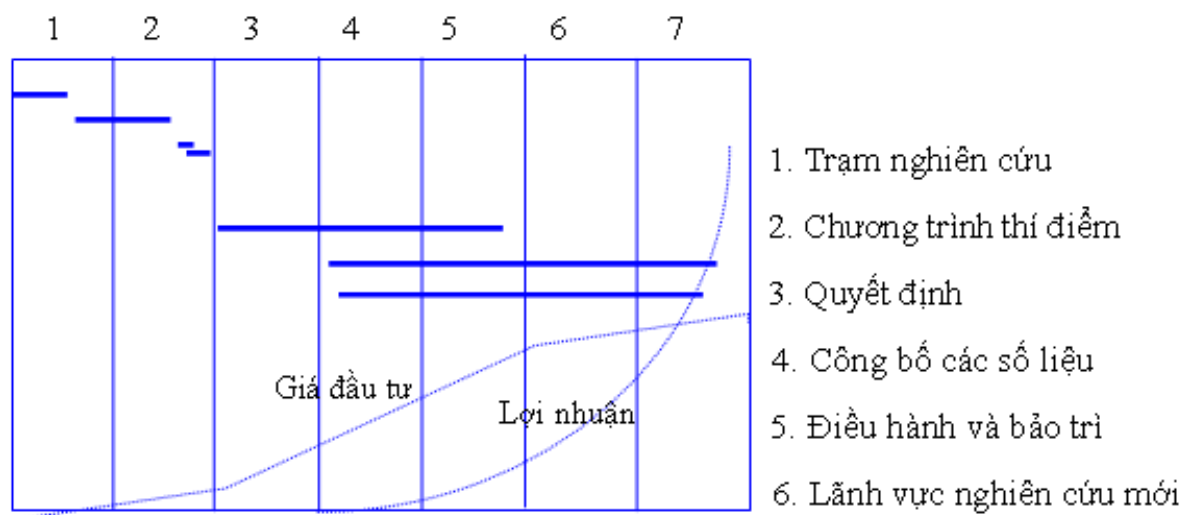
Trước khi bắt đầu xây dựng hoặc đề xuất các chương trình phát triển, cần thiết phải có một giai đoạn tiền khả thi cho một kế hoạch mang tính chiến lược. Kế hoạch đó nên chứa các đặc điểm như sau:

- Các đề nghị cho việc xây dựng các chương trình phụ nếu cần
- Mô tả các hoạt động sắp tới
- Sắp xếp các công việc nào cần thiết phải được xử lý bằng máy tính
- Xây dựng các giả định cần thiết để chứng minh các quyết định và lợi nhuận cũng như các đo lường về mặt kỹ thuật của các tổ chức cơ quan được yêu cầu

- Lập kế hoạch và xây dựng lịch làm việc cho các hoạt động khác nhau.
- Đề xuất các thành viên cần thiết cho việc thực hiện chương trình
- Xây dựng kế hoạch đầu tư và kinh phí cho chương trình
- Cần thiết phải tổ chức các khoá huấn luyện GIS, cùng với sự tham gia và hướng dẫn của các chuyên gia về GIS
- Tổ chức lại cơ cấu của các cơ quan
- Lập kế hoạch sử dụng kinh phí
- Tổ chức kế hoạch thực hiện các chương trình

Các kỹ thuật bao hàm trong GIS đang được phát triển một cách nhanh chóng và liên tục. Tuy nhiên người sử dụng GIS có khả năng cải tiến cũng như phát triển các ứng dụng mới của nó ở bất kỳ thời gian nào. Do đó, kế hoạch phát triển mang tính chiến lược nên mang tính linh động đủ để cho phép việc thực hiện được hiệu quả với các điều kiện đã nêu ở trên.

Việc sắp xếp các trình tự trong kế hoạch dựa vào kết quả của việc phân tích lợi nhuận, trong đó việc đầu tư cho GIS được kiểm chứng với việc đầu tư hiện có cũng như sắp tới ở một vài đơn vị nào đó của tổ chức. Thêm vào đó, các tổ chức hay cơ quan pháp nhân cần thiết phải xem xét và phân tích nhiệm vụ chính của từng đơn vị của họ.



Hình 10.1: Một điển hình cho sự khởi đầu GIS lý tưởng (Nguồn : Tor Bernhardsen, 1992)

Do đó một kế hoạch mang tính chiến lược nên đề ra và làm sáng tỏ các vấn đề sau:

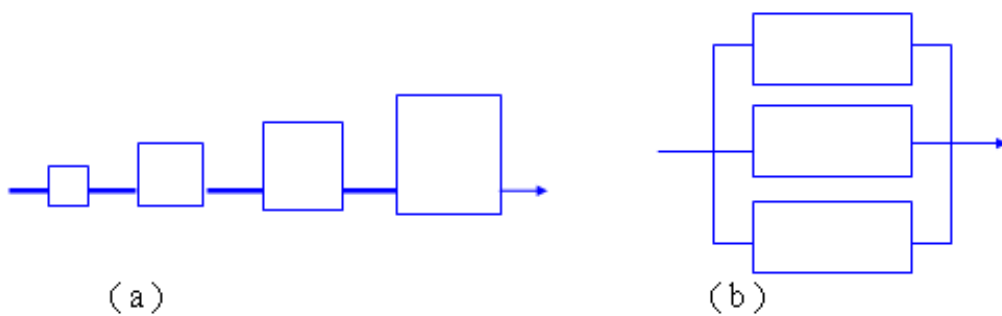
- Các công việc nào hay dữ liệu nào cần thiết phải được xử lý với GIS?

Nguồn số liệu hiện có, các đường dây thông tin, sự phân tích các sản phẩm và lợi nhuận là cơ sở cho việc hình dung ra những gì cần thiết phải được xử lý bằng GIS. Các hình thức đó bao gồm:

- Các công việc nào sẽ được thay bởi GIS?
- Các hồ sơ hoặc dữ liệu nào sẽ được thay thế bởi các dữ liệu GIS?
- Các nguồn số liệu sắp tới sẽ là những gì, bao nhiêu?

Việc mô tả các nguồn số liệu sẽ có sắp tới phải liên quan đến các mục đích cần thiết phải được giải quyết. Việc so sánh nguồn số liệu sẽ có với nguồn số liệu đang có sẽ giúp chỉ ra các công việc nào cần thiết giải quyết và các công việc nào phải được thêm vào cũng như các công việc nào nên được kiểm chứng.

Đối với các cơ quan hay tổ chức chính phủ, việc thay đổi toàn bộ nguồn số liệu do nguyên nhân của việc thay đổi cách quản lý thông thường với cách quản lý bằng GIS, chúng có thể được xem như là sự thay đổi từ hình thức liên tục đến hình thức song song. Các tiến trình công việc được trình bày qua hình sau đây:

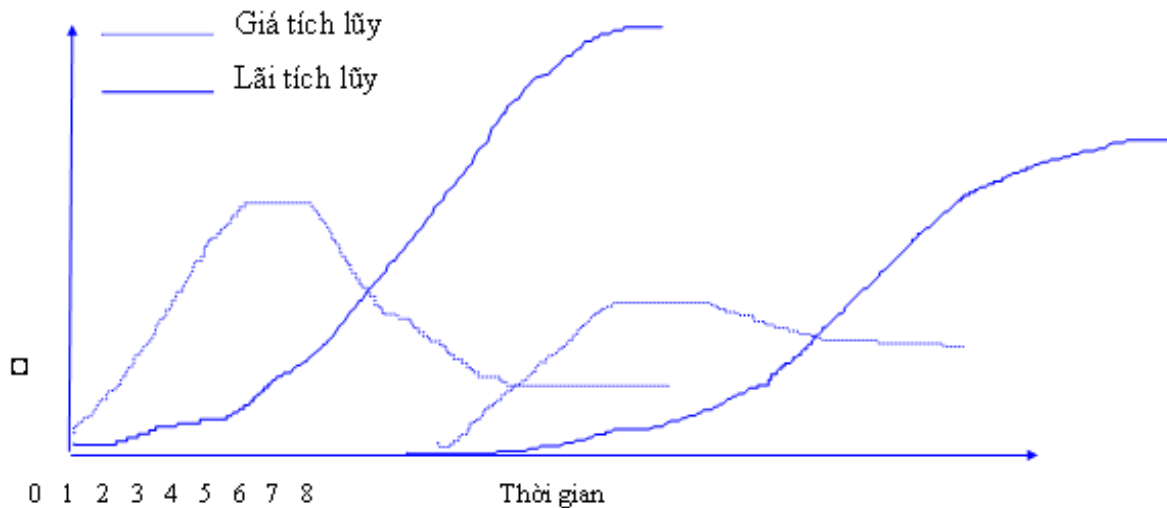


Hình 10.2: Sự thay đổi tiến trình thực hiện các công việc từ liên tục (a) đến song song (b) khi đưa các trang thiết bị GIS vào sử dụng (Nguồn : Tor Bernhardsen, 1992)

- Khi nào sẽ bắt đầu thực hiện chương trình sử dụng GIS?

Nhìn chung, sự khởi đầu cho việc sử dụng và ứng dụng GIS càng trễ thì kinh phí của việc thu thập và xử lý số liệu cũng như thời gian sẽ cao và dài hơn trước khi lợi nhuận được thu hồi. Tóm lại, các yếu tố ảnh hưởng đến việc quyết định khi nào thì bắt đầu việc sử dụng và ứng dụng GIS được trình bày như sau:

- Các kỹ thuật có liên quan thì hiện hữu nhưng chúng ta không biết bằng cách nào để khai thác nó
- Việc trì hoãn sẽ hoãn lại việc thu hồi lợi nhuận và từ đó lợi nhuận thu được từ các chương trình ít đi
- Tất cả các kỹ thuật cũng như tiến bộ hiện nay đều được diễn tiến liên tục do đó sẽ không có một thời gian tốt nhất cho việc khởi sự GIS.



Hình 10.3: Việc bắt đầu đưa vào sử dụng các trang thiết bị GIS sớm và trễ. (Nguồn : Tor Bernhardsen, 1992)

(Nếu bắt đầu sớm thì kết quả tỉ số lợi nhuận sẽ cao hơn bắt đầu trễ)

- Mức độ đầu tư cho chương trình là bao nhiêu?

Việc chọn lựa một quyết định được thực hiện trong việc đầu tư thường gặp những trở ngại giữa việc đầu tư các kỹ thuật mới khi so sánh giữa những chi phí thông thường và việc đầu tư cho các trang thiết bị mắc tiền hơn. Do đó, một số hướng dẫn cho các mức độ đầu tư có thể được tóm tắt như sau:

- Các suy nghĩ thì to lớn nhưng khi bắt đầu cần thiết phải thận trọng
- Xây dựng các đầu tư lớn sau khi thực hiện các chương trình thử nghiệm thí điểm
- Trong việc thay đổi từ cách quản lý và ứng dụng thông thường bằng cách quản lý GIS, việc đầu tư chính nên tập trung vào các phương tiện, trang thiết bị cho giai đoạn khởi đầu, từ đó chương trình sẽ mang lại nhiều lợi nhuận hơn.
- Các số liệu thu thập được nên nhanh chóng chuyển đổi sang các số liệu số (digital form)
- Càng nhiều đề tài được thực hiện trên cùng một vùng càng gia tăng lợi ích cũng như lợi nhuận cho người sử dụng

- Phạm vi hoạt động của chương trình hay GIS là bao nhiêu?

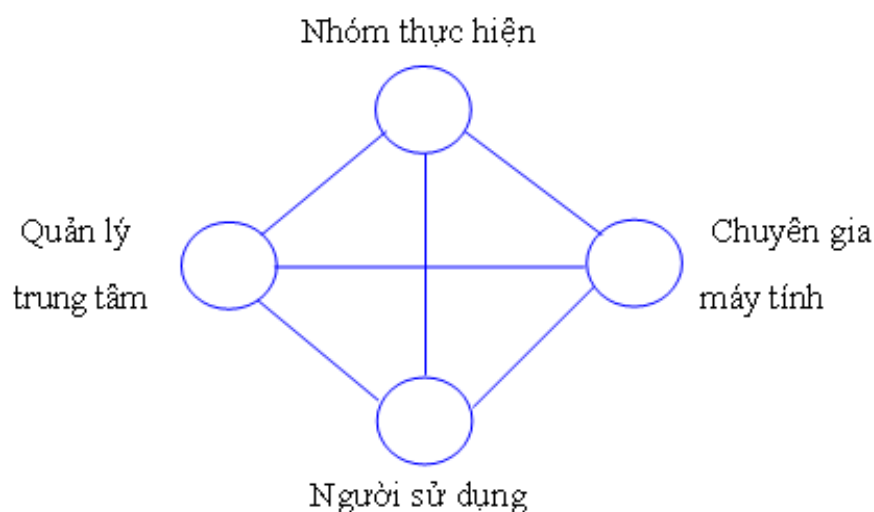
Về nguyên tắc, một quyết định cơ bản để thực hiện một chương trình trong một phạm vi nào đó của GIS có thể thu thập các số liệu số (digital) cho một phạm vi rộng của các vùng hay các thông tin đầy đủ cho một số vùng nào đó của một hoạt động ở mức độ chi tiết và nó sẽ thay đổi rất nhanh. Do đó, diễn tiến của vấn đề kinh tế đối với tỉ số lợi nhuận là một thông số quan trọng nhất. Tóm lại, một hướng dẫn cho một phạm vi hoạt động của GIS có thể được tóm lược như sau:

- Ít nhất phải có 1 đề tài bao trùm cho một vùng hoặc một khu vực cần thiết nào đó phải được chuyển đổi trước khi lợi nhuận được thu hồi.
- Khi đã chuyển sang GIS thì tất cả các số liệu xử lý theo cách trước đây nên được chuyển toàn bộ sang hình thức của GIS, tránh việc sử dụng hoặc xử lý đồng thời số liệu theo cách xử lý thông thường và theo cách của GIS.
- Việc tổ chức bộ phận quản lý cho một phạm vi hay cho một vùng nào đó nhất thiết phải được ưu tiên chú trọng, trong khi đó các tổ chức điều hành của một chương trình về cơ bản ít cần thiết hơn.
- Sự sắp xếp tổ chức đơn vị cơ quan như thế nào?

Hiệu quả của việc khai thác và khám phá các kỹ thuật mới trong một đơn vị hay tổ chức thường thay đổi tùy vào các chuỗi công việc mà nó ảnh hưởng đến toàn bộ cơ cấu tổ chức. Về thực tế, việc thay đổi cơ cấu tổ chức có thể gặp nhiều khó khăn vì cơ cấu tổ chức cũ đã được ổn định, do đó sẽ gặp khó khăn khi xây dựng một cơ cấu tổ chức mới. Việc thay đổi cơ cấu tổ chức cũng sẽ thay đổi các thành viên và các nhân sự có liên quan và việc thay đổi các thành viên luôn đưa đến mối liên hệ giữa con người, điều đó chắc sẽ đưa đến những khó khăn trong việc điều hành cũng như tiên đoán những sự việc sẽ xảy ra.

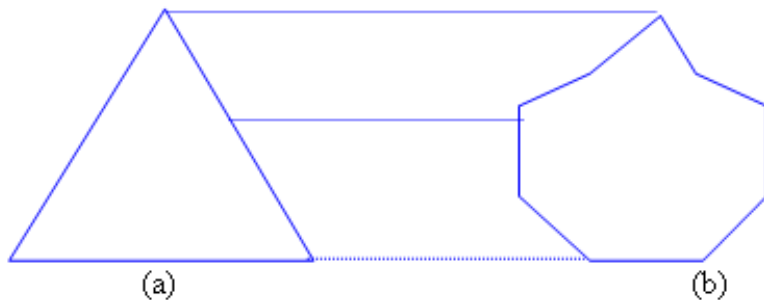
Về nguyên tắc, các khó khăn về mặt kỹ thuật có thể được giải quyết theo chiều thẳng tiến bằng việc mua sắm và cài đặt thêm các trang thiết bị mới, cũng như các phần mềm mới,... mặc dù nó sẽ đòi hỏi nhiều kinh

phí, tuy nhiên đối với các chương trình được lập kế hoạch và điều hành tốt, chi phí đó sẽ không đáng kể. Trong khi đó, việc thay thế các thành viên trong ban điều hành thì lại không theo chiều thẳng tiến và có thể phát sinh ra những khó khăn và trở ngại không theo ý muốn. Từ đó, việc tổ chức sắp xếp lại cơ cấu tổ chức thường đòi hỏi sự chú ý trong việc quản lý một cách thường xuyên và liên tục hơn là việc chú ý vào các khó khăn về mặt kỹ thuật. Một chương trình mang lại một kết quả tốt thường được thực hiện bằng cách quản lý chính diện, người ta thường tìm ra được những lợi nhuận của các kỹ thuật mới trong khi lại chống đối việc thay đổi cơ cấu tổ chức một cách mạnh mẽ. Do đó, việc tham gia của các chuyên gia là cần thiết trong việc thực hiện một chương trình GIS, nhưng họ lại ít có khả năng diễn đạt một chi tiết không liên quan trong chương trình. Thí dụ như các chuyên gia về máy tính thì thường luôn luôn tuân thủ các nguyên tắc và kỷ luật trong công việc của họ, trong khi đó họ lại thiếu một quan điểm bao quát được đòi hỏi có liên quan đến các khó khăn về mặt tổ chức.



Hình 10.4: Để đạt được thành công, một kỹ thuật mới phải được chấp nhận ở tất cả các cấp trong một tổ chức (Nguồn : Tor Bernhardsen, 1992)





Hình 10.5: Sự thay đổi trong quan điểm khi đưa ra một kỹ thuật mới trong một tổ chức (Nguồn : Tor Bernhardsen, 1992)

Như đã thảo luận ở trên, cần thiết phải có sự tham gia của các chuyên gia về GIS, nếu không tiến trình thực hiện có thể sẽ bị đình trệ hoặc công việc sẽ được thực hiện theo các chiều hướng khác không đúng với mục đích đã đề ra. Do đó việc thành lập và duy trì các thành viên chuyên gia nên được đặt ở vị trí ưu tiên. Việc quản lý nhân sự cho một giai đoạn chuyển giao các kỹ thuật mới nên bao gồm các yếu tố sau:

- Xác định các yếu tố ảnh hưởng đến nhân sự
- Lập kế hoạch cho việc đào tạo, hoặc bổ sung kiến thức và sắp xếp lại vị trí các nhân sự
- Nhấn mạnh rõ các công việc họ phải làm cũng như trách nhiệm của họ đối với công việc đó
- Các công nhân hoặc người làm thuê cũng được đưa vào trong kế hoạch
- Xác định rõ từng vị trí cũng như lương bổng cho từng thành viên
- Lập kế hoạch cho sự xoay vòng các công việc
- Diễn đạt và trình bày cơ cấu tổ chức phù hợp và đơn giản nhất, bao gồm:
  - Các thành viên chính thức trong ban quản lý
  - Các thành viên không chính thức được phép thay thế các thành viên chính thức trong trường hợp cần thiết
  - Phân chia cụ thể các công việc cho các thành viên
  - Nhấn mạnh rõ mối liên hệ giữa ban điều hành và bộ phận thực hiện, cũng như giữa và trong các tổ chức

- Nên giữ lại những điều thuận lợi cũng như sức mạnh của tổ chức
- Chứng nhận các thành viên cũng như các công việc của họ một cách thích hợp
- Tạo ra một môi trường với các công việc thử thách và đòi hỏi nhiều sự tập trung cũng như suy nghĩ để cải tiến kỹ thuật của họ

Do đó, các yếu tố về mặt tổ chức quan trọng nhất có thể được tóm lược như sau:

- Các khó khăn về mặt tổ chức thường nhiều hơn các khó khăn về mặt kỹ thuật
- Việc tổ chức hay sắp xếp lại nên được xem xét kỹ lưỡng
- Việc giới thiệu GIS ảnh hưởng đến những thay đổi trong chu trình hiện có của sự trao đổi thông tin giữa các đơn vị cũng như trong đơn vị của một tổ chức
- Các chu trình của công việc được thay đổi sẽ điều khiển việc thay đổi tổ chức
- Ít nhất là 1/4 thành viên của một tổ chức có thể sẽ được chuyển sang các công việc khác
- Việc hợp tác điều hành phải được lập ra một cách cụ thể
- Giai đoạn khởi đầu của sự lắp đặt các trang thiết bị cho GIS có thể được tổ chức như là một đề án. Các thay đổi của tổ chức nên được thử nghiệm trước khi đi đến việc thay đổi toàn bộ
- Việc thay đổi cơ cấu tổ chức về mặt lâu dài có thể được thực hiện sau giai đoạn thử nghiệm các trang thiết bị mới của GIS
- Kinh phí và hướng sử dụng kinh phí như thế nào?

Một chương trình lắp đặt các hệ thống của GIS nên có một kinh phí mà nó bao gồm cả sự phân bổ các hoạt động trong khoảng thời gian của chương trình và cho phép nó phải được kiểm chứng.

Các cơ quan thuộc về chính phủ thường phải được quyết định nên đầu tư kinh phí vào các trang thiết bị mới dựa vào nguồn kinh phí hiện có hoặc trích từ quỹ sử dụng theo tỉ lệ của thị trường.

- Hướng tổ chức thực hiện cũng như sắp xếp các thành viên và nhiệm vụ của họ trong chương trình như thế nào?

Về mặt thực tế, các chương trình thiết lập hệ thống GIS thường được thực hiện dưới sự góp ý và quyết định của các thành viên trong chương trình. Vì thế, các yêu cầu về thành phần các thành viên phải được xem xét một cách cẩn thận, và các thành viên khi được chỉ định phải sẵn sàng làm việc một cách có hiệu quả cho chương trình ở bất kỳ thời gian hay điều kiện nào. Về mặt nguyên tắc, điều này thường đòi hỏi các thành viên phải được bảo đảm an tâm về các mặt từ đó họ sẽ đầu tư hết thời gian cho công việc mà họ phải đảm trách đối với chương trình, tương tự cũng nên được áp dụng cho các cá nhân được thuê để phục vụ cho chương trình.

Một tổ chức mạnh cũng như được hoạt động độc lập thì sẽ dễ dàng thành công đối với bất kỳ việc thiết lập cũng như ứng dụng của GIS. Một chương trình về việc thiết lập hệ thống GIS không cần thiết phải được lập một cách vĩnh viễn, nó có thể được hoàn tất và giải tán khi các trang thiết bị GIS đã được đưa vào hoạt động. GIS, máy vi tính và các chuyên viên thì thường dễ dàng tìm kiếm hơn là phải đào tạo lại các nhân viên mới. Tất cả các chương trình nên cần những người hăng hái, say sưa với công việc cũng như những người có thể giải quyết được những vấn đề mà chương trình đặt ra.

Việc thuê mướn cũng như hợp đồng các chuyên gia từ bên ngoài để vào cùng thực hiện một chương trình thì thường có lợi hơn trong chiến lược của chương trình đó, họ không nhất thiết phải nằm trong ban điều hành cũng như ban tổ chức, mà nên sử dụng họ như những người có thể giải quyết những khó khăn và vướng mắc trong các lãnh vực cũng như các bất hoà hơn là những thành viên của tổ chức.

## Tổ chức cơ sở dữ liệu trong GIS

Chức năng của hệ thống thông tin địa lý là để cải thiện khả năng người sử dụng để đánh giá đưa đến sự quyết định trong nghiên cứu, qui hoạch và quản lý. Để sắp xếp cho một số hệ thống thông tin, người sử dụng cần phải được cung cấp dữ liệu một cách đầy đủ và hữu hiệu, điều này đạt được bởi phương pháp của hệ thống quản lý dữ liệu (DBMS). Một DBMS có thể được định nghĩa như sau: Một sự liên kết các dữ liệu đã lưu trữ cùng với nhau mà không gây một trở ngại hoặc việc làm dư thừa không cần thiết nhằm giúp ích cho chương trình được gia tăng khả năng sử dụng lên gấp bội; dữ liệu được lưu trữ để chúng là chương trình độc lập mà dữ liệu được sử dụng một cách phổ biến, và việc điều khiển trong việc thêm dữ liệu mới, hoặc sửa đổi và khôi phục dữ kiện hiện có bên trong hệ thống dữ liệu. Dữ liệu được kết cấu như thế để cung cấp một nền tảng cho việc phát triển sau này "(Martin, 1977)

## KHÁI NIỆM CƠ BẢN VỀ HỆ CƠ SỞ DỮ LIỆU

Để dễ dàng cho việc giải thích các khái niệm trước hết hệ thống bản vẽ máy bay bằng máy tính. Dữ liệu lưu trữ trong máy tính bao gồm thông tin về hành khách, chuyến bay, đường ..v.v.. Mọi thông tin về mối quan hệ này được biểu diễn trong máy thông qua việc đặt chỗ của khách hàng. Vậy làm thế nào để biểu diễn được dữ liệu đó và để đảm bảo cho hành khách đi đúng chuyến?

- Dữ liệu nêu trên được lưu trong máy theo một qui định nào đó và được gọi là cơ sở dữ liệu (CSDL, tiếng Anh là Database).
- Phần chương trình để có thể xử lý, thay đổi dữ liệu này là hệ quản trị cơ sở dữ liệu (HQTCSDL, Database Management System).

Theo nghĩa này HQTCSDL có nhiệm vụ rất quan trọng như là một bộ diễn dịch (Interpreter) với ngôn ngữ bậc cao nhằm giúp người sử dụng có thể dùng được hệ thống mà ít nhiều không cần quan tâm đến thuật toán chi tiết hoặc biểu diễn dữ liệu trong máy.

# CÁC LOẠI THÔNG TIN TRONG HỆ THỐNG THÔNG TIN ĐỊA LÝ

Như trên đã giới thiệu dữ liệu trong hệ thống thông tin địa lý bao gồm dữ liệu địa lý và dữ liệu thuộc tính:

\* Dữ liệu địa lý: bao gồm các thể loại

- Ảnh hàng không vũ trụ
- Bản đồ trực ảnh (orthophotomap)
- Bản đồ nền địa hình lập từ ảnh hàng không - vũ trụ
- Bản đồ địa hình lập từ số liệu đo đạc mặt đất
- Bản đồ địa chính
- Bản đồ địa lý tổng hợp từ các loại bản đồ địa hình.

Các loại ảnh và bản đồ nói trên đều ở dạng số và lưu lại dưới dạng vector hoặc raster hỗn hợp raster-vector. Các dữ liệu địa lý dưới dạng vector được phân lớp thông tin yêu cầu của việc tổ chức các thông tin. Thông thường người ta hay phân lớp theo tính chất thông tin: lớp địa hình, lớp thủy văn, lớp đường giao thông, lớp dân cư, lớp thức phụ, lớp giới hành chính v.v.. Trong nhiều trường hợp để quản lý sâu hơn, người ta sẽ phân lớp chuyên biệt hơn như trong lớp thủy văn được phân thành các lớp con: các lớp sông lớn, sông nhỏ, lớp biển, lớp ao hồ, v.v...

Các thông tin ở dạng raster là các thông tin nguồn và các thông tin hỗ trợ, không gian quản lý như một đối tượng địa lý. Các thông tin ở dạng vector tham gia trực tiếp quản lý và được định nghĩa như những đối tượng địa lý. Các đối tượng này thể hiện ở 3 dạng: điểm, đường và vùng hoặc miền. Mỗi đối tượng đều có thuộc tính hình học riêng như kích thước, miền vị trí. Vấn đề được đặt ra là tổ chức lưu trữ và hiển thị các thông tin vector như thế nào để thỏa mãn các yêu cầu sau:

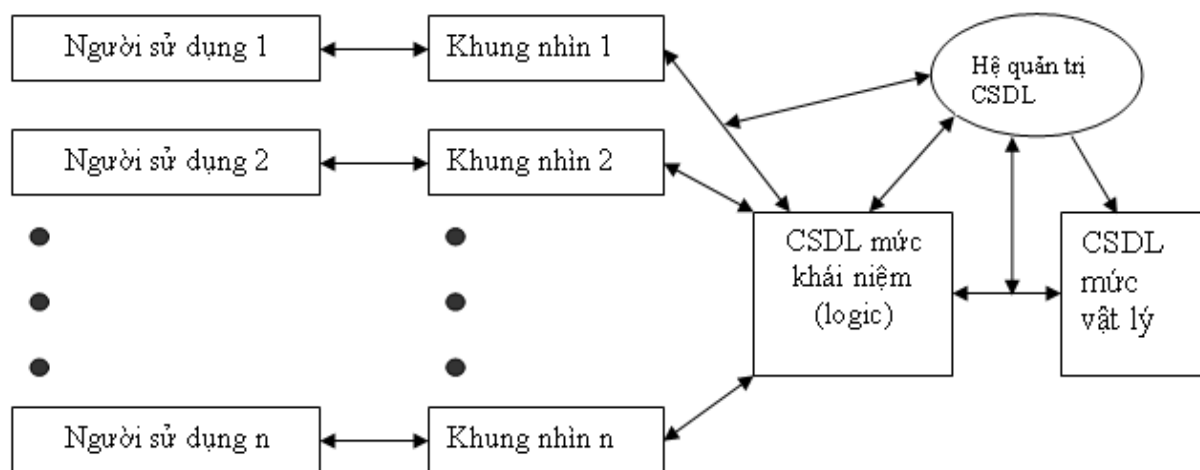
- Thể hiện đầy đủ các thông tin cần thiết
- Độ dư và độ thừa nhỏ nhất
- Truy cập thông tin nhanh

- Cập nhật thông tin dễ dàng và không sai sót (xoá bỏ thông tin không cần thiết, bổ sung thông tin mới, chỉnh lý các thông tin đã lạc hậu)
- Thuận lợi cho việc hiển thị thông tin

\* Dữ liệu thuộc tính (Attribute): là các thông tin giải thích cho các hiện tượng địa lý gắn liền với hiện tượng địa lý. Các thông tin này được lưu trữ dữ liệu thông thường. Vấn đề đặt ra là phải tìm mối quan hệ giữa thông tin địa lý và thông tin thuộc tính. Từ thông tin ta có thể tìm ra được các thông tin kia trong cơ sở dữ liệu.

## KIẾN TRÚC MỘT HỆ CƠ SỞ DỮ LIỆU

Một CSDL được phân thành các mức khác nhau. Ở đây có thể xem như chỉ có một CSDL đơn giản và có một hệ phần mềm QTCSDL.

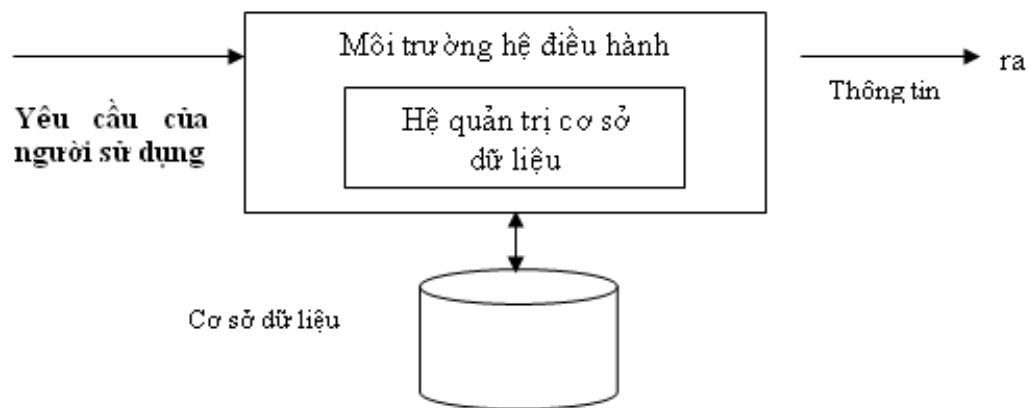


Hình 6.1: Cấu trúc hệ Cơ sở dữ liệu

### Sự trừu tượng hoá dữ liệu

- CSDL vật lý (mức vật lý) là các tệp dữ liệu theo một cấu trúc nào đó được lưu trên các thiết bị nhớ thứ cấp (như đĩa từ, băng từ, ...)

- CSDL mức khái niệm là một sự biểu diễn trừu tượng của CSDL vật lý (còn có thể nói tương đương: CSDL mức vật lý là sự cài đặt cụ thể của CSDL mức khái niệm).
- Các khung nhìn (view) là cách nhìn, là quan niệm của từng người sử dụng đối với CSDL mức khái niệm. Sự khác nhau giữa khung nhìn và mức khái niệm thực chất là không lớn.



Hình 6.2: Sơ đồ tổng quát hệ cơ sở dữ liệu

## Thể hiện và lược đồ của CSDL

### Thể hiện của CSDL (INSTANCE)

Khi CSDL đã được thiết kế, thường người ta quan tâm tới “bộ khung” hay còn gọi là “mẫu” của CSDL. Dữ liệu có trong CSDL gọi là thể hiện của CSDL, mặc dù khi dữ liệu thay đổi trong một chu kỳ thời gian nào đó thì “bộ khung” của CSDL vẫn không thay đổi.

CSDL luôn thay đổi mỗi khi thông tin được thêm vào hay bị xóa đi. Tập hợp các thông tin lưu trữ trong CSDL tại một thời điểm nào đó được gọi là một thể hiện của CSDL.

### **Lược đồ của CSDL (Scheme)**

Thiết kế tổng quan của CSDL được gọi là lược đồ (hay sơ đồ) của CSDL. Lược đồ của CSDL ít khi bị thay đổi. Trong một ngôn ngữ lập trình, nó tương ứng với các tập định nghĩa của các kiểu dữ liệu (kiểu mẫu tin, kiểu bảng, ...)

Thường “ bộ khung” nêu trên bao gồm một số danh mục hoặc chỉ tiêu hoặc một số kiểu của các thực thể trong CSDL. Giữa các thực thể có thể có mối quan hệ nào đó với nhau. Ở đây sử dụng thuật ngữ “ lược đồ” để thay thế cho khái niệm “ bộ khung”.

Lược đồ khái niệm là bộ khung của CSDL mức vật lý, khung nhìn được gọi là lược đồ con (Subscheme).

### **Lược đồ khái niệm và mô hình dữ liệu**

Lược đồ khái niệm là sự biểu diễn thế giới thực bằng một loại ngôn ngữ phù hợp, hệ QTCSDL cung cấp ngôn ngữ định nghĩa dữ liệu cho lược đồ con (subscheme data definition language) để xác định lược đồ khái niệm. Đây là ngôn ngữ bậc cao có khả năng mô tả lược đồ khái niệm bằng cách biểu diễn của mô hình dữ liệu. Ví dụ mô hình dữ liệu phù hợp là một đồ thị có hướng (mô hình mạng - Network model), trong đó các đỉnh biểu diễn mối quan hệ giữa các thực thể (như hành khách, chuyến bay), các cạnh của đồ thị biểu diễn mối quan hệ giữa các thực thể (như xác định đội bay cho mỗi chuyến bay).

### **Các mô hình của CSDL**

- Để dễ dàng minh họa và phân biệt giữa mô hình phân cấp và mô hình lưới, xem xét ví dụ sau đây:
- Cho một bản đồ đơn giản gồm 2 đa giác được xác định bởi tập hợp các đường thẳng trong có được một đường chung của 2 đa giác. Mỗi đường



thẳng được xác định bởi tập các đường thẳng trong đó có một đường chung của hai đa giác. Mỗi đường thẳng được xác định bởi các cặp tọa độ

Hình 6.3: Bản đồ A

Đa giác I gồm 4 cạnh a,b,c,d, với 4 đỉnh 1,2,3,4.

Đa giác II gồm 4 cạnh c,e,f,g với các đỉnh 3,4,5,6

#### **Mô hình phân cấp (HIERACHICAL)**

Mô hình dữ liệu là một cây, trong đó các nút biểu diễn các tập thực thể, giữa các nút con và nút cha được liên hệ theo một mối quan hệ xác định. Điểm nổi bật trong các thủ tục truy xuất đến một đối tượng trong mô hình phân cấp là đường dẫn đi từ gốc đến phần tử cần xét trong cây phân cấp.

Hình 6.4: Biểu diễn bản đồ A bằng mô hình phân cấp (Nguồn : Phạm Trọng Mạnh, Phạm Vọng Thành, 1999)

Mô hình phân cấp khá phù hợp với những hình thức tổ chức phân cấp trong xã hội. Thường gặp trong các hệ thống máy tính là mô hình quản lý thư mục

#### **Mô hình lưới (Network Model)**

Mô hình dữ liệu kiểu lưới là mô hình cho phép dùng một mô hình đồ thị trực tiếp và đơn giản cho dữ liệu.

Để dễ dàng minh họa và phân biệt giữa mô hình phân cấp và mô hình lưới, xem xét ví dụ sau đây:

Cho một bản đồ A đơn giản gồm 2 đa giác I và II được xác định bởi tập hợp các đường thẳng trong đó có được một đường chung của 2 đa giác.

Mỗi đường thẳng được xác định bởi các cặp toạ độ.

Hình 6.5 Biểu diễn bản đồ A bằng mô hình lưới (Nguồn : Phạm Trọng Mạnh, Phạm Vọng Thành, 1999)

Mô hình lưới và mô hình phân cấp nói chung là khá bất tiện cho lưu trữ và khai thác xử lý bởi vì toạ độ các điểm, một số cạnh phải lưu trữ nhiều lần (như ví dụ trên các cạnh c phải lưu trữ 2 lần) v.v.. gây nên sự dư thừa dữ liệu. Ngoài ra, hệ thống còn phải cần lưu trữ một số lớn các con trỏ móc nối gây nên phức tạp trong quá trình cập nhật, biến đổi dữ liệu, đặc biệt khi thêm bớt một cạnh hoặc một đỉnh nào đó.

#### **Mô hình quan hệ (Relational Model)**

Mô hình này dựa trên cơ sở khái niệm lý thuyết tập hợp của các quan hệ, tức là tập các K - bộ với K cố định.

Thuận lợi của mô hình quan hệ là được hình thức hoá toán học chặt chẽ do đó các xử lý, thao tác với dữ liệu là dễ dàng, có tính độc lập dữ liệu cao. Cấu trúc dữ liệu đơn giản mềm dẻo trong xử lý và dễ dàng cho người sử dụng. Đặc biệt các phép tính cập nhật dữ liệu cho mô hình quan hệ nói chung là ít phức tạp hơn nhiều so với các mô hình khác.

Một cách đơn giản hơn có thể hiểu mối quan hệ là một bảng 2 chiều tập độc lập, trong đó mỗi cột (trường) là một thuộc tính, mỗi hàng (bộ) là một đối tượng. Trong thí dụ trên, có cấu trúc các quan hệ (bảng) như sau:

Hình 6.5 Biểu diễn bản đồ A bằng mô hình quan hệ (Nguồn : Phạm Trọng Mạnh, Phạm Vọng Thành, 1999)

Trong 3 loại mô hình nêu trên thì mô hình quan hệ có nhiều ưu điểm và được nhiều người quan tâm hơn cả. Bởi lẽ, mô hình dữ liệu quan hệ có tính độc lập rất cao, lại dễ dàng sử dụng. Điều quan trọng hơn cả, mô hình quan hệ được hình thức hoá toán học tốt, do đó được nghiên cứu,

phát triển và cho được nhiều kết quả lý thuyết cũng như ứng dụng trong thực tiễn.

Trên cơ sở mô hình dữ liệu quan hệ, đến nay đã phát triển thêm một số loại mô hình khác nhằm mô tả và thể hiện thế giới thực một cách chính xác và phù hợp hơn như mô hình quan hệ thực thể (Entity Relationship model), mô hình dữ liệu hướng đối tượng (Object Oriented Model)

## **Tính độc lập dữ liệu (Data independence)**

**Sự phụ thuộc dữ liệu của các ứng dụng hiện nay**

Một hệ ứng dụng được gọi là phụ thuộc dữ liệu khi không thể thay đổi được cấu trúc lưu trữ hoặc chiến lược truy xuất mà không ảnh hưởng đến chương trình ứng dụng.

Lúc đó, các khó khăn gây nên bởi cấu trúc mặt ghép nối của các chương trình xử lý tập tin, chứ không liên quan gì đến vấn đề mà hệ ứng dụng viết ra để giải quyết.

**Yêu cầu của các hệ ứng dụng**

- Các hệ ứng dụng khác nhau cần có những cái nhìn khác nhau đối với những dữ liệu khác nhau.
- Người quản trị cơ sở dữ liệu phải có khả năng thay đổi cấu trúc lưu trữ hoặc chiến lược truy xuất để đáp ứng các yêu cầu thay đổi mà không cần phải sửa đổi các chương trình ứng dụng hiện có.

**Định nghĩa tính độc lập dữ liệu**

Tính độc lập dữ liệu là tính bất biến của các hệ ứng dụng đối với các thay đổi trong cấu trúc lưu trữ và chiến lược truy xuất.

#### **Phân loại tính độc lập dữ liệu**

\* Độc lập dữ liệu ở mức vật lý:

Theo hình trên, từ khung nhìn, tới CSDL khái niệm và CSDL vật lý cho thấy có hai mức “độc lập tư liệu”. Thứ nhất: Lược đồ có thể thay đổi do người quản trị CSDL mà không cần thay đổi lược đồ con. Việc tổ chức lại CSDL vật lý (thay đổi các tổ chức, cấu trúc dữ liệu trên các thiết bị nhớ thứ cấp) có thể làm thay đổi hiệu quả tính toán của các chương trình ứng dụng nhưng không đòi hỏi phải viết lại các chương trình đó. Tính độc lập này gọi là Độc lập dữ liệu mức vật lý.

\* Độc lập dữ liệu ở mức logic:

Mối quan hệ giữa các khung hình và lược đồ khái niệm cho thêm một loại độc lập, gọi là độc lập dữ liệu logic. Khi sử dụng một CSDL, có thể cần thiết để thay đổi lược đồ khái niệm như thêm thông tin về các loại khác nhau của các thực thể hoặc bớt xoá các thông tin về các thực thể đang tồn tại trong CSDL. Việc thay đổi lược đồ khái niệm không làm ảnh hưởng tới các lược đồ con đang tồn tại, do đó không cần thiết phải thay đổi các chương trình ứng dụng.

## **HỆ QUẢN TRỊ CƠ SỞ DỮ LIỆU CỦA GIS**

### **Giới thiệu**

Hệ quản trị cơ sở dữ liệu là hệ thống quản lý, lưu trữ, bảo trì toàn bộ cơ sở dữ liệu. Hệ quản trị cơ sở dữ liệu cũng cung cấp các công cụ cho phép người dùng hỏi đáp, tra cứu và tác động vào cơ sở dữ liệu. Cơ sở dữ liệu cho một hệ thống GIS bao gồm 2 cơ sở dữ liệu thành phần chính là :

- Cơ sở dữ liệu địa lý (không gian)
- Cơ sở dữ liệu thuộc tính (phi không gian)

Trong hệ thống GIS, hệ quản trị cơ sở dữ liệu GIS được xây dựng bao gồm 2 hệ quản trị cơ sở dữ liệu riêng cho từng phần hoặc xây dựng một hệ quản trị cơ sở dữ liệu chung cho cả hai cơ sở dữ liệu con kể trên. Thông thường hệ quản trị cơ sở dữ liệu GIS được xây dựng bao gồm 3 hệ quản trị cơ sở dữ liệu con:

- Hệ quản trị cơ sở dữ liệu cho cơ sở dữ liệu địa lý.
- Hệ quản trị cơ sở dữ liệu quan hệ ở mức tra cứu, hỏi đáp. Hệ này được tích hợp cùng với hệ quản trị cơ sở dữ liệu địa lý cho phép người ta dùng truy nhập dữ liệu địa lý và dữ liệu thuộc tính đồng thời. Tuy nhiên, hệ quản trị cơ sở dữ liệu này cho thao tác trên cơ sở dữ liệu thuộc tính bị hạn chế.
- Hệ quản trị cơ sở dữ liệu thuộc tính. Thông thường các hệ thống GIS đều lấy một hệ quản trị cơ sở dữ liệu quan hệ hiện có để quản trị và thực hiện các bài toán trên dữ liệu thuộc tính mà không liên quan đến dữ liệu không gian. Ví dụ: FOX, MS SQL, ORACLE.

Về hệ quản trị cơ sở dữ liệu quản lý cơ sở dữ liệu thuộc tính, chúng ta đã xem xét chi tiết trong phần “Hệ thống cơ sở dữ liệu“. Vì vậy, ở đây chúng ta chỉ đi sâu vào hệ quản trị cơ sở dữ liệu cho dữ liệu không gian.

Hệ quản trị cơ sở dữ liệu cho dữ liệu không gian bao gồm các hệ thống con sau:

- Hệ thống nhập bản đồ
- Hệ thống hiển thị bản đồ
- Hệ thống tra cứu, hỏi đáp cơ sở dữ liệu
- Hệ thống phân tích địa lý
- Hệ thống phân tích thống kê
- Hệ thống đầu ra

## **Hệ quản trị cơ sở dữ liệu GIS**

### **Hệ thống nhập bản đồ**

Hệ thống cung cấp các công cụ để số hoá các đối tượng trên bản đồ. Hiện nay tồn tại hai phương pháp để chuyển bản đồ giấy thành bản đồ số:

- Số hoá bản đồ: dùng bàn vẽ (digitizer) đi lại các đối tượng bản đồ trên giấy. Chức năng có thể có trong bản thân hệ thống hoặc dùng 1 phần mềm khác số hoá, sau đó nhập vào kết quả số hoá bởi phần mềm đó.
- Vector hoá bản đồ: Bản đồ được quét vào thành dạng file ảnh (scanning) sau đó chuyển sang dạng vector (vectorizing). Quá trình vectorizing có thể thực hiện thủ công qua số hoá trên màn hình (head up digitizing) hoặc dùng phần mềm chuyển tự động/bán tự động từ ảnh sang vector.

Một trong những nguồn dữ liệu quan trọng là dữ liệu được nhập từ các hệ thống khác. Vì vậy hệ thống nhập bản đồ phải có chức năng nhập (import) các dạng (format) dữ liệu khác nhau.

### **Hệ thống hiển thị bản đồ**

Hệ thống cung cấp các khả năng hiển thị bản đồ trên màn hình cho người sử dụng xem. Hiện nay chức năng hiển thị bản đồ đều có khả năng cung cấp cách nhìn 3 chiều (3D). Bản đồ sẽ được thể hiện sinh động, trực quan hơn.

### **Tra cứu, hỏi đáp cơ sở dữ liệu**

Hệ thống cung cấp các công cụ cho người sử dụng tra cứu, hỏi đáp, lấy các thông tin cần thiết trong cơ sở dữ liệu ra. Hệ thống này ảnh hưởng trực tiếp đến tính hiệu quả của hệ thống. Bởi vì phần lớn, người dùng chỉ có yêu cầu tra cứu thông tin. Hệ thống tra cứu phải mềm dẻo, dễ sử

dụng, thời gian truy cập dữ liệu nhanh. Hệ thống cho phép tra cứu trên cả hai dữ liệu: dữ liệu không gian và dữ liệu thuộc tính

#### **Hệ thống xử lý, phân tích địa lý**

Đây là hệ thống thể hiện rõ nhất sức mạnh của GIS. Hệ thống cung cấp các công cụ cho phép người dùng xử lý, phân tích dạng dữ liệu không gian. Từ đó, chúng ta có thể sản sinh ra các thông tin mới (thông tin dẫn suất)

#### **Hệ thống phân tích thống kê**

Hệ thống cung cấp các công cụ thống kê trên dữ liệu không gian cũng như dữ liệu thuộc tính. Tuy nhiên các phép phân tích thống kê trên dữ liệu không gian khác biệt so với một số phép phân tích thống kê thông thường trên dữ liệu phi không gian.

#### **Hệ thống in ấn bản đồ**

Hệ thống có nhiệm vụ in các bản đồ kết quả ra các thiết bị ra thông dụng như máy in (printer), máy vẽ (Plotter). Yêu cầu đối với hệ thống này là tương thích với nhiều loại thiết bị ngoại vì hiện có trên thị trường.

## CPU Structure and Functions

This module presents the processor organization. We will begin with a summary of processor organization. Registers, which form the internal memory of the processor, are then analyzed. We are then in a position to return to the discussion of the instruction cycle. A description of the instruction cycle and a common technique known as instruction pipelining completes our description. Understanding the organization of the CPU of a computer

### Processor Organization

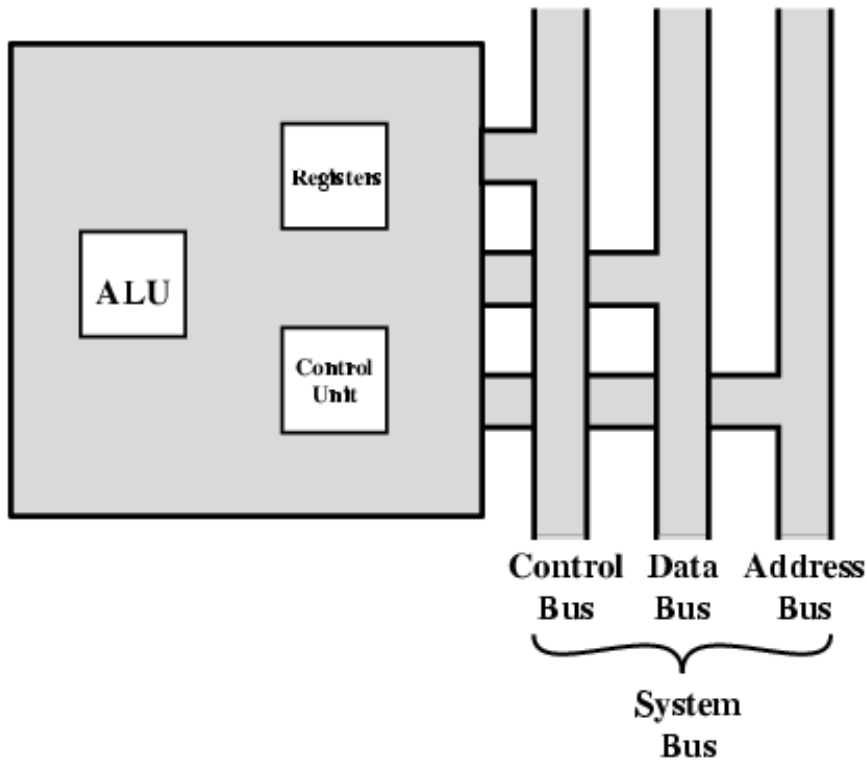
To understand the organization of the CPU, let us consider the requirements placed on the CPU, the things that it must do:

- **Fetch instruction:** The CPU reads an instruction from memory.
- **Interpret instruction:** The instruction is decoded to determine what action is required.
- **Fetch data:** The execution of an instruction may require reading data from memory or an I/O module.
- **Process data:** The execution of an instruction may require performing some arithmetic or logical operation on data.
- **Write data:** The results of an execution may require writing data to memory or an I/O module.

To do these things, it should be clear that the CPU needs to store some data temporarily. It must remember the location of the last instruction so that it can know where to get the next instruction. It needs to store instructions and data temporarily while an instruction is being executed. In other words, the CPU needs a small internal memory.

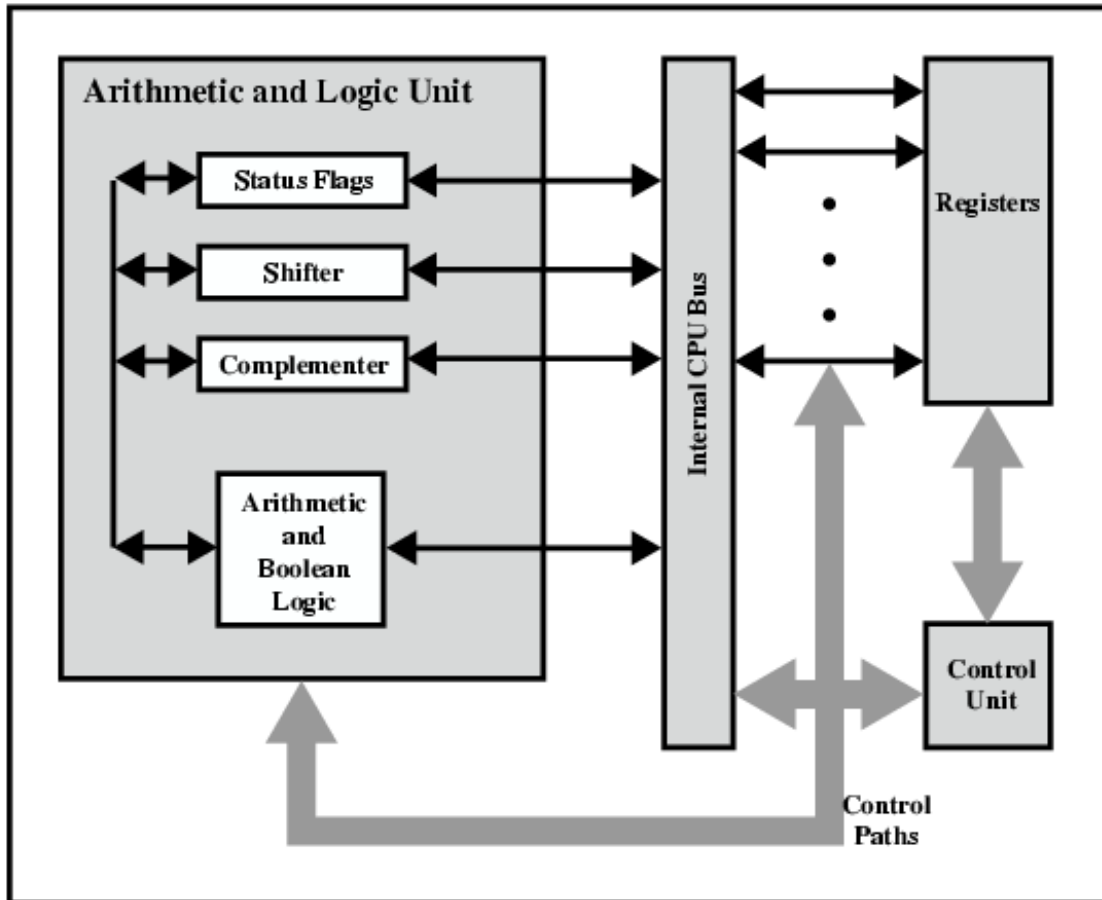
[\[link\]](#) is a simplified view of a CPU, indicating its connection to the rest of the system via the system bus. You will recall (Lecture 1) that the major components of the CPU are an **arithmetic and logic unit** (ALU) and a **control unit** (CU). The ALU does the actual computation or processing of data. The control unit controls the movement of data and instructions into and out of the CPU and controls the operation of the ALU. In addition, the figure shows a minimal internal memory, consisting of a set of storage locations, called **registers**.





The CPU with the System Bus

[\[link\]](#) is a slightly more detailed view of the CPU. The data transfer and logic control paths are indicated, including an element labeled **internal CPU-bus**. This element is needed to transfer data between the various registers and the ALU because the ALU in fact operates only on data in the internal CPU memory.



CPU Internal Structure

## Register organization

Within the CPU, there is a set of registers that function as a level of memory above main memory and cache in the hierarchy. The registers in the CPU perform two roles:

- **User-visible registers:** These enable the machine- or assembly-language programmer to minimize main memory references by optimizing use of registers.
- **Control and status registers:** These are used by the control unit to control the operation of the CPU and by privileged, operating system

programs to control the execution of programs.

There is not a clean separation of registers into these two categories. For example, on some machines the program counter is user visible (e.g., Pentium), but on many it is not (e.g., PowerPC). For purposes of the following discussion, however, we will use these categories.

## **User-Visible Registers**

A user-visible register is one that may be referenced by means of the machine language that the CPU executes. We can characterize these in the following categories:

- General purpose
- Data
- Address
- Condition codes

**General-purpose registers:** can be assigned to a variety of functions by the programmer. Sometimes their use within the instruction set is orthogonal to the operation. That is, any general-purpose register can contain the operand for any opcode. This provides true general-purpose register use. Often, however, there are restrictions. For example, there may be dedicated registers for floating-point and stack operations. In some cases, general-purpose registers can be used for addressing functions (e.g., register indirect, displacement). In other cases, there is a partial or clean separation between data registers and address registers.

**Data registers** may be used only to hold data and cannot be employed in the calculation of an operand address.

**Address registers** may themselves be somewhat general purpose, or they may be devoted to a particular addressing mode. Examples include the following:

- **Segment pointers:** In a machine with segmented addressing, a segment register holds the address of the base of the segment. There may be multiple registers: for example, one for the operating system and one for the current process.
- **Index registers:** These are used for indexed addressing and may be autoindexed.
- **Stack pointer:** If there is user-visible stack addressing, then typically the stack is in memory and there is a dedicated register that points to the top of the stack. This allows implicit addressing; that is, push, pop, and other stack instructions need not contain an explicit stack operand.

**Condition codes register** (also referred to as **flags**): Condition codes are bits set by the CPU hardware as the result of operations. For example, an arithmetic operation may produce a positive, negative, zero, or overflow result. In addition to the result itself being stored in a register or memory, a condition code is also set. The code may subsequently be tested as part of a conditional branch operation.

## Control and Status Registers

There are a variety of CPU registers that are employed to control the operation of the CPU. Most of these, on most machines, are not visible to the user. Some of them may be visible to machine instructions executed in a control or operating system mode.

Of course, different machines will have different register organizations and use different terminology. We list here a reasonably complete list of register types, with a brief description.

Four registers are essential to instruction execution:

- **Program counter (PC):** Contains the address of an instruction to be fetched.
- **Instruction register (IR):** Contains the instruction most recently fetched.

- **Memory address registers (MAR):** Contains the address of a location in memory.
- **Memory buffer register (MBR):** Contains a word of data to be written to memory or the word most recently read.

Typically, the CPU updates the PC after each instruction fetch so that the PC always points to the next instruction to be executed. A branch or skip instruction will also modify the contents of the PC. The fetched instruction is loaded into an IR, where the opcode and operand specifiers are analyzed. Data are exchanged with memory using the MAR and MBR. In a bus-organized system, the MAR connects directly to the address bus, and the MBR connects directly to the data bus. User-visible registers, in turn, exchange data with the MBR.

The four registers just mentioned are used for the movement of data between the CPU and memory. Within the CPU, data must be presented to the ALU for processing. The ALU may have direct access to the MBR and user-visible registers. Alternatively, there may be additional buffering registers at the boundary to the ALU: these registers serve as input and output registers for the ALU and exchange data with the MBR and user-visible registers.

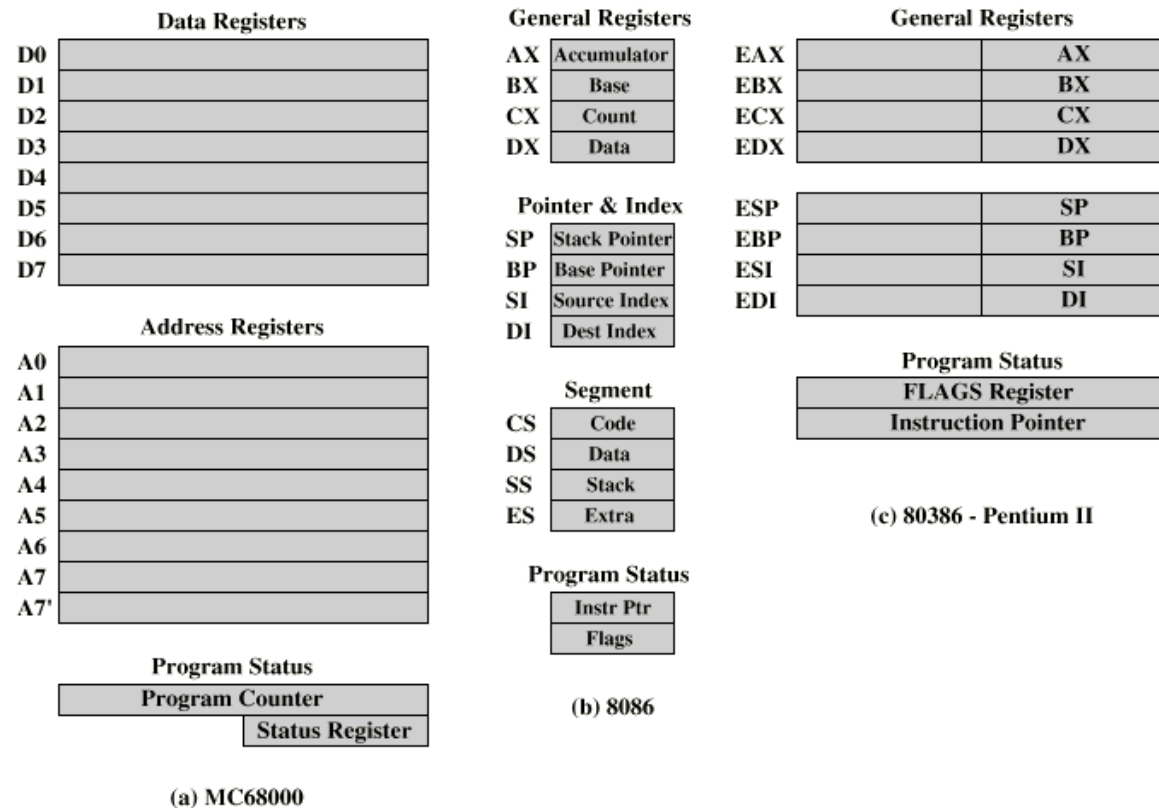
All CPU designs include a register or set of registers, often known as the **program status word (PSW)**, that contain status information. The PSW typically contains condition codes plus other status information. Common fields or flags include the following:

- **Sign:** Contains the sign bit of the result of the last arithmetic operation.
- **Zero:** Set when the result is 0.
- **Carry:** Set if an operation resulted in a carry (addition) into or borrow (subtraction) out of a high-order bit. Used for multiword arithmetic operations.
- **Equal:** Set if a logical compare result is equality.
- **Overflow:** Used to indicate arithmetic overflow
- **Interrupt enable/disable:** Used to enable or disable interrupts.
- **Supervisor:** Indicates whether the CPU is executing in supervisor or user mode. Certain privileged instructions can be executed only in

supervisor mode, and certain areas of memory can be accessed only in supervisor mode.

A number of other registers related to status and control might be found in a particular CPU design. In addition to the PSW, there may be a pointer to a block of memory containing additional status information (e.g., process control blocks).

Example Register Organizations:



Example of microprocessor registers organizations